Réalisation stochastique de signaux non stationnaires et identification sur un seul echantillon

Albert Benveniste, Jean-Jacques Fuchs

INRIA

CENTRE DE RENNES

IRISA

Institut National
de Recherche
en Informatique
et en Automatique

Domaine de Voluceau
Rocquencourt
BP 105
78153 Le Chesnay Cedex
France
Tél. 954 90 20

Rapports de Recherche

N° 128

RÉALISATION STOCHASTIQUE DE SIGNAUX NON STATIONNAIRES, ET IDENTIFICATION SUR UN SEUL ÉCHANTILLON

Albert BENVENISTE
Jean-Jacques FUCHS

Avril 1982

RÉALISATIONS STOCHASTIQUES DE SIGNAUX NON STATIONNAIRES, ET IDENTIFICATION SUR UN SEUL ÉCHANTILLON

Albert BENVENISTE (+)

Jean-Jacques FUCHS (+)

Publication Interne n° 158 - Mars 1982 - 33 pages

RESUME : On montre que la méthode des Variables Instrumentales, ainsi que l'algorithme de réalisation de Ho-Kalman, sont consistants pour l'identification sur un seul échantillon des caractéristiques modales (partie pôle) d'un processus Gaussien-Markovien excité par un bruit blanc non stationnaire.

ABSTRACT : Gauss-Markov processes excited by nonstationary noises are encountered in the modelling of vibrating systems. We prove that the classical Instrumental Variable method, as well as the Ho-Kalman realization algorithm, for identifying the pole part (modal characterisics) of the model, are consistent when used on a single sample of the (nonstationary) signal.

(+) Membre IEEE IRISA/INRIA

Secrêtariat de Rédaction : Melle F. Moinet -IRISA- Lab. Informatique
Université de Rennes - Campus de Beaulieu
35042 RENNES CEDEX

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE
(L.A. 227)
UNIVERSITÉ DE RENNES 1, I.N.S.A. DE RENNES

INSTITUT NATIONAL DE RECHERCHE
EN INFORMATIQUE ET EN AUTOMATIQUE
(LABORATOIRE DE RENNES)

I. INTRODUCTION AND MAIN RESULTS.

State space models of the form

$$(I.1)\quad \begin{cases} X_{t+1} = FX_t + V_{t+1} & \text{(F asymptotically stable)} \\ Y_t = HX_t \ , \end{cases}$$

where the observation (Y_t) is a vector process, and the excitation (V_t) is a (non-stationary) white noise with time-varying covariance matrix $\mathbb{E}(V_{t+1}V_{t+1}^T) = Q_t$, are often encountered in various applications. For example, vibrating systems are often subject to non stationary excitation noises ([1][2][3]), and may be described by such models. Note that the corresponding signal (Y_t) is non stationary, although the matrices H and F are constant.

The problem we are interested in is the following :

PROBLEM : Given a single sample $Y_0, Y_1, \dots, Y_S$ of the process (Y_t), how to identify (up to a change of basis in the state space) the matrices H and F ?

Note that it is not clear that this is possible, since no ergodicity property can be used for this purpose in view of the non stationary character of the observed signal (Y_t).

An answer to this problem would provide us with the modal characteristics of the signal (Y_t), namely the pairs $(\lambda, H\phi_\lambda)$ where λ is an eigenvalue of F, and ϕ_λ the corresponding eigenvector ; and it turns out ([1][2][3]) that this corresponds to the modal decomposition of the underlying vibrating structure in the case of vibration measurements.

The non stationary stochastic realization of signals of the form (1) is analysed in [4][5][6] when multiple data records are available ; this is an entirely different problem, since in this case the non stationary covariance function $\Gamma(i,j) = \mathbb{E}\, Y_i Y_j^T$ may be estimated by averaging upon the different records of the variables Y_i and Y_j ; such problems are encountered in econometrics where non stationary phenomena are periodically repeated. Unfortunately, this is not the case in our situation, where only one sample is available.

Let us analyse the problem from a heuristical viewpoint, by looking at scalar non stationary AR MA signal of the form

$$(\mathrm{I},2) \qquad y_t = \sum_{n=1}^{N} a_n y_{t-n} + \sum_{p=0}^{P} b_p(t)\, e_{t-p}$$

where (e_t) is a standartwhite noise and the MA coefficients $b_p(t)$ are slowly time varying. Then, for S large, let us decompose the sample $y_0, \ldots, y_S$ into blocks of the form $y_{kT}, y_{kT+1}, \ldots, y_{kT+T-1}$ for $k=0,\ldots,K-1$; which are small enough for the b_j's to be considered as constant, and large enough for the true parameters a_i being the solution of the system of equations:

$$(\mathrm{I}.3) \qquad \begin{aligned} \sum_{n-1}^{N} a_n \Gamma_{P+1-n}^{(k)} &\cong \Gamma_{P+1}^{(k)} \\ \sum_{n=1}^{N} a_n \Gamma_{P+N-n}^{(k)} &\cong \Gamma_{P+N}^{(k)} \end{aligned}$$

where

$$(\mathrm{I}.4) \qquad \Gamma_n^{(k)} = \frac{1}{T} \sum_{t=kT}^{kT+1-1} y_t y_{t-n}$$

is the empirical covariance function of (y_t) estimated on the k-th block, where the system is considered as stationary. In the classical stationary case, (I.3) corresponds to the classical Instrumental Variable method with delayed observations ([7]-[11]) which is known to be consistent, so that the relationships (I.3) are approximately valid in our case, by assuming (y_t) to be stationary inside each block. Now, combining (I.3) for various k gives

$$(\mathrm{I}.5) \qquad \left\{ \begin{aligned} \sum_{n=1}^{N} a_n \left(\frac{1}{K} \sum_{K=0}^{K-1} \Gamma_{P+1-r}^{(k)}\right) &\cong \frac{1}{K} \sum_{k=0}^{K-1} \Gamma_{P+1}^{(k)} \\ &\vdots \\ \sum_{r=1}^{N} a_n \left(\frac{1}{K} \sum_{k=0}^{K-1} \Gamma_{P+N-r}^{(k)}\right) &\cong \frac{1}{K} \sum_{k=0}^{K-1} \Gamma_{P+N}^{(k)} \end{aligned} \right.$$

But this shows that $(a_1,\dots,a_N)$ is the solution of the Instrumental Variable method applied to the whole sample $y_0,\dots,y_S$, without taking care of the non stationary character of the signal, since

$$(I.6) \qquad \frac{1}{K}\sum_{k=0}^{K-1} \Gamma_n^{(k)} = \frac{1}{S}\sum_{0}^{S-1} y_t\, y_{t-n}$$

is nothing but the single-sample empirical covariance function of (y_t).

The reason for this "consistency result" is the linear character of the instrumental Variable method. So it may be expected that any linear estimator should be consistent, including the Ho-Kalman realization algorithm for example. On the other hand, non linear estimates like Maximum likelihood are expected to be nonrobust in case of non stationary excitation.

Finally, note that the known consistency results for the Instrumentel Variable method require the asymptotic stationarity of the involved signals [7][8][9][10].

Now let us state the main results for the case of model (I.1). Let us introduce the following assumptions :

ASSUMPTIONS 1 (i) The pair (H,F) is observable

(ii) The covariance matrix Q_t is bounded uniformly in t.

(iii) There exists a matrix G_* such that the pair (F,G_*) is controllable, and such that, with probability 1, we have for S large :

$$(I.7) \qquad \frac{1}{A_S}\sum_{t=0}^{S-1} V_{t+1}\, V_{t+1}^T \geq G_* G_*^T ,$$

where

$$(I.8) \qquad A_S = \sum_{t=0}^{S-1} ||x_t||^2 ,$$

is assumed to satisfy $A_S \to +\infty$ w.p.1 as $S \to \infty$.

COMMENT : Condition (iii) is a kind of (strong) uniform controllability condition. The following result is proved in Appendix C :

LEMMA 1 : The following condition implies condition (iii) of assumption 1 : there exists G_* with

(I.9) $Q_t > G_* G_*^T$ for every t, and (F,G_*) controllable.

In fact, condition (iii) is much weaker than the present one, since it is easy to see that this condition allows changes in the geometry of the excitation matrix Q_t, which is not the case in the condition (I.9).

(I.1) Robustness of the Instrumental Variable method.

Now we are able to give the main results. Let us assume Y_t having values in $\mathbb{R}^d$, and X_t having values in $\mathbb{R}^p$ (i.e. the process is "of order p") : let n ($\leq p$) be the smallest index such that the observability matrix:

$$(\text{I.10}) \qquad \mathcal{O}_n = \begin{bmatrix} H \\ HF \\ \\ HF^{n-1} \end{bmatrix}$$

is of full rank ; and let $A_1, \ldots, A_n$ be a solution of

$$(\text{I.11}) \qquad HF^n = \sum_{i=1}^{n} A_i \; HF^{n-i} = [A_n, \ldots, A_1] \, \mathcal{O}_n ;$$

note that such a solution is generally non unique, unless the rows of $\mathcal{O}_n$ are independent. Then the pair

$$(\text{I.12}) \qquad \hat{H} = [I, 0 \ldots 0], \; \hat{F} = \begin{bmatrix} 0 & I & & \\ & \ddots & \ddots & \\ & & 0 & I \\ A_n & \text{—} & \text{—} & A_1 \end{bmatrix}$$

is a (non-minimal) realization of the pair (H,F), in the sense that the modal characteristics $(\lambda, H\phi_\lambda)$ of the system (I.1) belong to the set of the pairs (λ, Ψ_λ) where $\lambda \in \mathbb{C}$ and

$$(\lambda^n I - \sum_{i=1}^{n} \lambda^{n-i} A_i)\, \Psi_\lambda = 0. \tag{I.13}$$

Let us introduce the corresponding Instrumental Variable method. Set

$$\Gamma_n(S) = \sum_{t=0}^{S-n} Y_{t+n}\, Y_t^T \quad (n>0), \tag{I.14}$$

which is the empirical covariance function of the (non stationary) sample $Y_0,\dots Y_S$; reeall that in our situation, $\frac{1}{S}\Gamma_n(S)$ has no limit when $S\to\infty$ since no ergodic property is assumed. Consider the Hankel matrices

$$\mathbb{H}_{n,N}(S) = \begin{bmatrix} \Gamma_0(S) & \text{———} & \Gamma_n(S) & \text{———} & \Gamma_N(S) \\ | & \diagup & & \diagup & | \\ \Gamma_n(S) & \text{———} & & \Gamma_N(S) \;\text{—} & \Gamma_{N+n}(S) \end{bmatrix}$$

(I.15)

$$\mathbb{H}_n(S) \triangleq \mathbb{H}_{n,n}(S).$$

Then let us introduce the following estimation methods :

<u>Instrumental Variable (I.V.) method</u> : Solve the system

$$[-\hat{A}_n,\dots,-\hat{A}_1,\ I]\ \mathbb{H}_n(S) = 0 \tag{I.16}$$

and denote by $\hat{A}_1(S),\dots,\hat{A}_n(S)$ a solution.

Far Past Projection (FPP) method of order N : Solve in the least squares sense the system

$$(I.17) \qquad [-\hat{A}_n,\dots,-\hat{A}_1,\ I]\ \mathcal{H}_{n,N}(S) = 0$$

and denote by $\hat{A}_1^N(S),\dots,\hat{A}_n^N(S)$ a solution.

Now let us state the first main result :

THEOREM 1 : Let the assumptions 1 be in force. Then, there exists an integer N, depending only upon the bound given in assumption 1-i, and upon G_*, such that the FPP method of order N is consistent in the sense that

$$\lim_{S\to\infty} \sum_{i=1}^{n} (\hat{A}_i^N(S)-A_i)\ HF^{n-i} = 0 \text{ w.p.1},$$

where $A_1,\dots,A_n$ is an arbitrary solution of (I.11).

Recall that the IV method is consistent in the stationary case (Q_t=Q), which corresponds to N=n in the theorem 1.

Clearly, a drawback of this method lies in the fact that the reconstructed pair $(\hat{H},\hat{F})$ given in (I.12) is non minimal, or, equivalently, that the matrix coefficents $A_1,\dots,A_n$ are not unique. Let us investigate this point.

(I.2) Robustness of the Ho-Kalman algorithm [12][16]

Let us denote by $\mathcal{O}^{HO}$ any invertible square matrix obtained by extracting rows from $\mathcal{O}_n$, and let us denote by $Z.\mathcal{O}^{HO}$ the matrix obtained by extracting the same rows in the shifted observability matrix

$$(I.18) \qquad Z.\ \mathcal{O}_n \triangleq \begin{bmatrix} HF \\ | \\ HF^n \end{bmatrix} \ ;$$

we have

$$(I.19) \qquad Z.\mathcal{O}^{HO} = \mathcal{O}^{HO}F \ .$$

Hence, the pair $(\hat{H},\hat{F})$ defined by

$$(I.20) \quad H=\hat{H}\,\mathcal{O}^{HO} \ , \quad Z.\,\mathcal{O}^{HO} = \hat{F}\,\mathcal{O}^{HO}$$

is now a minimal realization of the pair (H,F), in the sense that they are identical up to a change of basis in the state space. These remarks lead to the following method.

HO-Kalman method : select in the Hankel matrix $\mathcal{H}_n(S)$ an invertible pxp matrix ($p=\dim X_t$), and denote by $\mathcal{H}^{HO}(S)$ the corresponding matrix ; let $Z.\mathcal{H}_n(S)$ be the shifted matrix, obtained by deleting the first p-block row in the matrix $\mathcal{H}_{n+1,n}(S)$, and denote by $Z.\mathcal{H}^{HO}(S)$ the matrix obtained in selecting the same rows and columns in $Z.\mathcal{H}_n(S)$ as before. Then solve for $\hat{H}(S)$ and $\hat{F}(S)$ the systems

$$(I.21) \qquad \begin{aligned} \mathcal{H}_{0,n}(S) &= \hat{H}(S)\ \mathcal{H}^{HO}(S) \\ Z.\mathcal{H}^{HO}(S) &= \hat{F}(S)\ \mathcal{H}^{HO}(S), \end{aligned}$$

thus obtaining estimates $\hat{H}(S)$ and $\hat{F}(S)$ for $\hat{H}$ and $\hat{F}$, which were defined in (I.20) .

Ho-Kalman method, with Far Past Projection of order N :

Proceed in the same way as before, but starting with $\mathcal{H}_{n,N}(S)$ and $Z.\mathcal{H}_{n,N}(S)$ respectively ; and denote by $\mathcal{H}_N^{HO}(S)$ the $p\times((N+1)d)$ matrices obtained by selecting p independent rows in $\mathcal{H}_{n,N}(S)$, and the same rows in $Z.\mathcal{H}_{n,N}(S)$, respectively.

Then solve in the least squares sense the systems

$$(I.22) \qquad \begin{aligned} \mathcal{H}_{0,N}(S) &= \hat{H}^N(S).\mathcal{H}_N^{HO}(S) \\ Z.\mathcal{H}_N^{HO}(S) &= \hat{F}^N(S).\ \mathcal{H}_N^{HO}(S), \end{aligned}$$

thus obtaining estimates $\hat{H}^N(S)$ and $\hat{F}^N(S)$ of $\hat{H}$ and $\hat{F}$.

Assume that the selection operation can be done independently of S (for S large), the "true matrices" $\hat{H}$ and $\hat{F}$ being obviously obtained in selecting the same rows in the observability matrix. Then we have again the following consistency result :

THEOREM 2. Let the assumptions 1 be in force. Then there exists N (depending upon the bound i), and G_*) such that the Ko-Kalman method with Far Past Projection of order N is consistent :

$$\hat{H}^N(S) \to \hat{H} \,,\; \hat{F}^N(S) \to \hat{F} \,,\; \text{w.p.1} \,. \tag{I.23}$$

Again, recall that the classical Ho-Kalman method is consistent in the stationary case ($Q_t = Q$).

Other classical realization algorithms ([13][14]) should also be robust in the same sense. However, analysing in the same non stationary context the robustness of efficient algorithms based upon the Singular Value Decomposition of the Hankel matrix $\mathcal{H}_\infty(S)$ ([18][19][20][21]) seems to be a much harder task, we have not investigated at the moment.

(I.3) The case of models with measurement noise.

In some systems, measurement noise should also be taken into account thus giving the model

$$\begin{cases} X_{t+1} = FX_t + V_{t+1}, & \text{COV } V_{t+1} = Q_t \\ Y_{t+1} = HX_t + W_{t+1}, & \text{COV } W_{t+1} = R_t, \end{cases} \tag{I.24}$$

where the noises (V_t) and (W_t) are assumed to be independent. Let us modify the assumptions 1 in the following way :

ASSUMPTIONS 2 (i) The pair (H,F) is observable.

(ii) The covariance matrices Q_t and R_t are bounded uniformly in t

(iii) There exists a constant matrix G_*, such that the pair (F, FG_*) is controllable, and such that, for S large, the conditions (I.7) and (I.8) hold w.p.1

Note that R_t may be singular. Set

$$\text{(I.25)} \qquad \bar{H}_{n,N}(S) = \begin{bmatrix} \Gamma_1(S) & \text{—} & \Gamma_{n+1}(S) & \text{———} & \Gamma_{N+1}(S) \\ | & \diagup & & \diagup & | \\ \Gamma_{n+1}(S) & \text{———} & \Gamma_{N+1}(S) & \text{—} & \Gamma_{N+n+1}(S) \end{bmatrix},$$

and denote by $\bar{A}_1^N(S),\ldots,\bar{A}_n^N(S)$ the least squares solution to the system

$$\text{(I.26)} \quad [-\bar{A}_n^N(S),\ldots,-\bar{A}_1^N(S),I]\ \bar{H}_{n,N}(S) = 0$$

which is the equivalent in this case of the FPP method of order N. Then the following theorem holds :

THEOREM 3 : Let the assumptions 2 be in force. Then there exists N (depending on the bounds in assumption 2-i) and on G_*) such that the FPP method of order N is consistent, in the sense that

$$\text{(I.27)} \qquad \lim_{S\to\infty} \sum_{i=1}^{r} (\bar{A}_i^N(S)-A_i)\ HF^{n-i}=0 \text{ w.p. } 1,$$

where the A_i's satisfy (I.11).

Of course, the corresponding results for the Ho-Kalman method should also be stated in this case, but we have dropped them for the sake of shortness.

(I.4) Organization of the sequel of the paper

The sequel of the paper is devoted to the proofs of the theorems.

The proof will proceed as follows. The empirical Hankel matrix (I.15) will be decomposed into two parts : in the first one only the empirical covariances of the state X_t and the noise V_t appear, whereas the second one contains the contribution of the empirical cross-correlation of X_t and V_{t+i} for $i>0$. A martingale argument is used for showing that the later part is neglectible with respect to A_S ; on the other hand, lower bounds are needed on the singular values of the first part, for showing that this first part is not neglectible with respect to A_S.

Thus the section II is entirely devoted to the derivation of bounds for the singular values of Hankel matrices associated to stationary processes. The section III is devoted to the end of the proof.

Finally, extensive experimental results are reported in [1][2][3], where the robustness of the FPP method is shown in the nonstationary case.

II. UPPER AND LOWER BOUNDS FOR THE SINGULAR VALUES OF HANKEL MATRICES OF GAUSS-MARKOV PROCESSES.

Throughout this section, only stationary Gauss-Markov processes are involved.

II.1. Gauss-Markov processes without measurement noise.

Let us consider the following stationary Gauss-Markov process :

$$\text{(II.1)} \qquad \begin{cases} X_{t+1} = FX_t + GV_{t+1} \\ Y_t = HX_t \end{cases}$$

where the matrix F is assumed to be asymptotically stable (hence (Y_t) is regular), and (V_t) is a standard white noise with identity covariance matrix the matrices H, F, G are respectively dxp, pxp, pxn. Assume also the representation (II.1) to be minimal, i.e. the pairs (H,F) and (F,G) are respectively observable and controllable.

Let us consider the two following (infinite dimensional) Hankel matrices.

$$\text{(II.2)} \qquad \mathcal{H}(H,F,G) = \begin{bmatrix} HG & HFG & HF^2G & \cdots & HF^kG & \cdots \\ HFG & HF^2G & & \iddots & & \\ HF^2G & & \iddots & & & \\ \vdots & \iddots & & & & \\ HF^kG & & & & & \\ \vdots & & & & & \end{bmatrix}$$

and

$$(\text{II.3}) \qquad \mathbb{H}(Y_t) = \begin{bmatrix} \Gamma_0 & \Gamma_1 & \text{---} & \Gamma_k & \text{---} \\ \Gamma_1 & & & & \\ | & & & & \\ \Gamma_k & & & & \\ | & & & & \end{bmatrix}, \quad \Gamma_k = \mathbb{E}\, Y_{t+k} Y_t^T .$$

By assumption of minimality of the representation (II.1), both Hankel matrices $\mathbb{H}(H,F,G)$ and $\mathbb{H}(Y_t)$ are of rank p ; the question we are interested in is to evaluate <u>how far is the process</u> (Y_t) <u>from being of rank</u> p-1 ? Let be

$$(\text{II.4}) \qquad \begin{aligned} &\sigma_{max}(H,F,G) \triangleq \sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_p \triangleq \sigma_{min}(H,F,G) \\ &\sigma_{min}(Y_t) \triangleq \sigma_1(Y_t) \geq \ldots \geq \sigma_p(Y_t) \triangleq \sigma_{min}(Y_t) \end{aligned}$$

the first p singular values (†) of the matrices (H,F,G) and (Y_t) respectively (the other ones are zero). And set

$$(\text{II.5}) \qquad \begin{aligned} S_{max}(Y_t) &= \max_{0<\theta<2\pi} (\lambda_{max}(S(e^{i\theta}))) \\ S_{min}(Y_t) &= \min_{0<\theta<2\pi} (\lambda_{min}(S(e^{i\theta}))) \end{aligned}$$

(where $\lambda_{max}(A)$ and $\lambda_{min}(A)$ denote respectively the greatest and lowest eigenvalue of the matrix A) where

$$(\text{II.6}) \qquad S(Z) = H(I-z F)^{-1} GG^T (I-z^{-1}F^T)^{-1} HT$$

is the spectrum of the process (Y_t). Finally, let

$$(\text{II.7}) \qquad \begin{aligned} &\mathbb{C}(F,G) = (G, FG, F^2G, \ldots) \\ &\mathbb{C}(F,PH^T) = (PH^T, FPH^T, \ldots), \text{ where } P = \mathbb{C}(F,G)\, \mathbb{C}^T(F,G) \\ &\mathbb{O}(H,F) = (H^T, F^TH^T, F^{2T}H^T, \ldots)^T \end{aligned}$$

(†) See Appendix A for definition and properties of the singular values of a matrix.

be the controlability matrices of the pairs (F,G) and (F,PH^T), and the observability matrix of the pair (H,F) ; and let be respectively

$$\text{(II.8)}\quad \begin{aligned} &\sigma_{max}(\mathcal{C}) \triangleq \sigma_1(\mathcal{C}) \geqslant \ldots \geqslant \sigma_p(\mathcal{C}) \triangleq \sigma_{min}(\mathcal{C}) \\ &\sigma_{max}(\mathcal{C}^P) \triangleq \sigma_1(\mathcal{C}^P) \geqslant \ldots \geqslant \sigma_p(\mathcal{C}^P) \triangleq \sigma_{min}(\mathcal{C}^P) \\ &\sigma max(\mathcal{O}) \triangleq \sigma_1(\mathcal{O}) \geqslant \ldots \geqslant \sigma_p(\mathcal{O}) \triangleq \sigma_{min}(\mathcal{O}) \end{aligned}$$

their corresponding singular values.

Let us assume that the process (Y_t) is <u>regular</u> and <u>of full rank</u> ([15][22]), namely that

$$\text{(II.9)}\quad \text{dét } S(e^{i\theta}) \neq 0 \text{ for } 0>\theta>2\pi.$$

Then, we have the following result :

THEOREM 4 : <u>Under assumption</u> (II.9), <u>the following inequalities hold</u> :

$$\text{(II-10-a)}\quad \begin{aligned} &\sigma_{max}(Y_t) \leqslant \sigma_{max}(H,F,G)\ (S_{max}(Y_t))^{1/2} \\ &\sigma_{max}(H,F,G) \leqslant \sigma_{max}(\mathcal{O})\quad \sigma_{max}(\mathcal{C}), \end{aligned}$$

<u>and</u>

$$\text{(II-10-b)}\quad \begin{aligned} &\sigma_{min}(Y_t) \geqslant \sigma_{min}(H,F,G)\ (S_{min}(Y_t))^{1/2} > 0 \\ &\sigma_{min}(H,F,G) \geqslant \sigma_{min}(\mathcal{O})\quad \sigma_{min}(\mathcal{C}), \\ &\sigma_{min}(\mathcal{C}^P) \geqslant \sigma_{min}(\mathcal{C})\ (S_{min}(Y_t))^{1/2} \end{aligned}$$

Before giving the proof of the theorem 4, let us give a corollary and a refined version of them.

COROLLARY 1 : Let us consider the following Lyapunov equation :

$$Q = FQF^T + PH^THP \tag{II.11}$$

where $P>0$ is assumed to be such that

$$P - FPF^T = GG^T(\geq 0). \tag{II.12}$$

Then, we have the following inequalities

$$\begin{aligned} &\lambda_{max}(Q) \leq \lambda_{max}(P)\ S_{max}\ (Y_t) \\ &\lambda_{min}(Q) \geq \lambda_{min}(P)\ S_{min}(Y_t), \end{aligned} \tag{II.13}$$

where S_{min} and S_{max} are defined in (II.5) and (II.6).

To our knowledge, this is a new result in the area of matrix Lyapunov equations.

Now, for an arbitrary set $\{i_1,\ldots,i_m\}$, let us denote by $\sigma_{max}^{i_1,\ldots,i_m}(Y_t)$ the greatest singular value of the matrix obtained by selecting in $\mathbb{H}(Y_t)$ the rows $i_1,\ldots,i_m$; and let us use the same notation for the other matrices. We have the following refinement of theorem 4 :

THEOREM 5 : The following inequalities hold

$$\begin{aligned} &\sigma_{max}^{i_1,\ldots,i_m}(Y_t) \leq \sigma_{max}^{i_1,\ldots,i_m}(H,F,G)\ [S_{max}(Y_t)]^{1/2} \\ &\sigma_{max}^{i_1,\ldots,i_m}(H,F,G) \leq \sigma_{max}^{i_1,\ldots,i_m}(\mathbb{O})\ \sigma_{max}(\mathbb{C}) \end{aligned} \tag{II.14}$$

together with the corresponding inequalities with "$\ldots_{min}$" instead of "$\ldots_{max}$".

PROOFS : The proof of theorem 4 rests upon the following factorization of the Hankel matrix $\mathcal{H}(Y_t)$:

$$\text{(II.15)} \qquad \mathcal{H}(Y_t) \;=\; \mathcal{H}(H,F,G) \,.\, \mathcal{R}^{T/2} \;=\; \mathcal{O}(H,F) \,.\, \mathcal{C}(F,G) \,.\, \mathcal{R}^{T/2} \;,$$

where

$$\text{(II.16)} \qquad \mathcal{R}^{1/2} = \begin{bmatrix} HG & HFG & HF^2G \ldots & \\ 0 & HG & HFG & \ddots \\ 0 & 0 & HG & \ddots \\ \vdots & & \ddots & \ddots \end{bmatrix}$$

is a square root of the covariance matrix of (Y_t), namely

$$\text{(II.17)} \qquad \mathcal{R}^{1/2}\,\mathcal{R}^{T/2} = \mathcal{R} \triangleq \begin{bmatrix} \Gamma_o & \Gamma_1 & \ddots \\ \Gamma_1^T & \Gamma_o & \ddots \\ & \ddots & \ddots \end{bmatrix} .$$

Then, taking into account the fact that de process is regular and of full rank (condition (II.9)), we get $\mathcal{R} > 0$; then we get (II.10 - a) and (II.10 - b) thanks to (A - 4,5,6,11) of appendix A, using the fact that the lowest and greatest eigenvalues of $\mathcal{R}$ are equal to $S_{min}(Y_t)$ and $S_{max}(Y_t)$ respectively ([23]). This finishes the proof of theorem 4.

For the corollary, note that, under assumption (II.12), the matrix Q is given by

$$\text{(II.18)} \qquad Q = Q^{1/2}\, Q^{T/2} \quad , \quad Q^{1/2} = \mathcal{C}(F,G)\, \mathcal{R}^{T/2}.$$

Finally the theorem 5 proved in the same way as the theorem 4, by keeping only the desired rows of the matrices $\mathcal{H}(Y_t)$, $\mathcal{H}(H,F,G)$, and $\mathcal{O}(H,F)$.

WARNING : There is no corresponding lower bound for the singular values of the matrix obtained by selecting columns (instead of rows) in the Hankel matrix $\mathbb{H}(Y_t)$, except in the case of a scalar signal (d = 1). The reason is that the corresponding factorization of $\mathbb{H}(Y_t)$ would involve the extraction of columns in $\mathbb{R}^{T/2}$, thus obtaining a (non - invertible) rectangular matrix, what is not allowed by (A.11). This is not surprising since the construction of a basis of the state space of (Y_t) involves the extraction of linearly independant rows of the Hankel matrix ([15]) which is not equivalent to the extraction of columns, except in the scalar case.

(II.2) Regular Gauss-Markov Process of full rank.

Now, let us give the corresponding results in the classical case of Gauss-Markov processes with minimal representation ([15])

$$\text{(II.19)} \quad \begin{cases} X_{t+1} = F\, X_t + V_{t+1} \\ Y_{t+1} = H\, X_t + W_{t+1} \end{cases},$$

where we are interested in getting bounds for the singular values of the shifted Hankel matrix

$$\text{(II.20)} \quad \bar{\mathbb{H}}(Y_t) \triangleq \begin{bmatrix} \Gamma_1 & \Gamma_2 & \text{---} & \Gamma_k & \text{---} \\ \Gamma_2 & & & & \\ | & & & & \\ \Gamma_k & & & & \\ | & & & & \end{bmatrix}$$

Let us focus our attention on the bounds which are independent of the choice of the state space representation ; thus we can choose the state space representation corresponding to the strong factorization of the spectrum of (Y_t), which is characterized by ([15] [16]) :

$$\text{(II.21)} \quad \mathbb{E}\begin{pmatrix} V_t \\ W_t \end{pmatrix} (V_t^T \; W_t^T) = \begin{pmatrix} K \\ I \end{pmatrix} R \, (K^T \; I) ,$$

where R is the covariance matrix of the innovation process of (Y_t) and K is the Kalman gain, i.e. the unique solution of the wellknown Algebraic Riccati Equation such that F - KH be asymptotically stable ([15] [16]). Then we have the following result, assuming that (Y_t) is <u>regular</u> and <u>of full rank</u> :

<u>THEOREM 6</u> : <u>The following inequalities hold for the process</u> (II.19) :

$$\text{(II.22)} \quad \begin{array}{lll} \bar{\sigma}_{max}(Y_t) & \leq & \sigma_{max}(H,F,KR^{1/2})\ (S_{max}(Y_t))^{1/2} \\ \bar{\sigma}_{min}(Y_t) & \geq & \sigma_{min}(H,F,KR^{1/2})\ (S_{min}(Y_t))^{1/2} > 0 \end{array}$$

Here, $\bar{\sigma}_i(Y_t)$ denote the singular values of $\bar{\mathcal{H}}(Y_t)$ (cf. (II.20)) ordered as in (II.4), whereas $\sigma_i(H,F,KR^{1/2})$ denote the singular values of the Hankel matrix $\mathcal{H}(H,F,G)$ with $G = KR^{1/2}$. Note that $(H(I - ZF)^{-1} K + I)\ R^{1/2}$ is the strong factorization of the spectrum of (Y_t), so that the bounds in (II.22) are indeed canonical, i.e. independent of the choice of the state space realization of (Y_t) .

<u>PROOF</u> (sketchy) : Here, the relevant factorization of the matrix $\bar{\mathcal{H}}(Y_t)$ is the following one :

$$\text{(II.23)} \quad \bar{\mathcal{H}}(Y_t) = \mathcal{H}(H,F,KR^{1/2})\ .\ \mathcal{R}^{T/2} ,$$

where

$$\text{(II.24)} \quad \mathcal{R}^{1/2} = \begin{bmatrix} R^{1/2} & HKR^{1/2} & HFKR^{1/2} & HF^2KR^{1/2} & \dots \\ 0 & R^{1/2} & HKR^{1/2} & HFKR^{1/2} & \ddots \\ \vdots & & \ddots & \ddots & \ddots \end{bmatrix}$$

is again a square root of the covariance matrix R of the process (Y_t) (cf. ((II.17) ; hence the theorem.

Note that the singular values σ_{min} and σ_{max} in (II.22) do not depend upon the choice of the square root of R.

REMARK : Finally, let us mention that, for state space representations of the form (II.19), such that (V_t) be independent of (w_t), with covariance

$$\text{(II.25)} \qquad \mathbb{E}\; V_t V_t^T = G\,G^T \quad ,$$

the following inequalities hold

$$\text{(II.26)} \qquad \begin{aligned} \bar{\sigma}_{max}(Y_t) &\leq \sigma_{max}(H,F,FG)\;(\tilde{S}_{max})^{1/2} \\ \bar{\sigma}_{min}(Y_t) &\geq \sigma_{min}(H,F,FG)\;(\tilde{S}_{min})^{1/2} \end{aligned} \quad ,$$

where $\sigma(H,F,FG)$ are defined as before, where as $\tilde{S}_{max/min}$ are defined as in (II.5) but using

$$\text{(II.27)} \qquad \tilde{S}(z) \triangleq H(I - zF)^{-1}\; G\,G^T(I - z^{-1}F^T)^{-1}\; H^T = S(z) - \mathrm{cov}(w_t) \quad ,$$

instead of the whole spectrum of (Y_t).

Of course, the later bounds (II.26) are coarser than the preceeding ones given in theorem 6, but they will be useful in the sequel ; moreover these bounds do not require that $\mathrm{cov}(w_t)$ be $>$ 0.

where

$$\text{(III.4)} \qquad \mathcal{F}_t = \sigma\{V_o, \ldots, V_t\}.$$

Let us compute the predictable quadratic variation $\langle M,M\rangle_S$ of this martingale ([25]) :

$$\langle M,M\rangle_S - \langle M,M\rangle_{S-1} \triangleq \mathbb{E}\left((M_S - M_{S-1})^2 \,\middle|\, \mathcal{F}_{S+i-1}\right)^{(*)}$$

$$= \mathbb{E}\left(H_k F^m V_{S+i} \; X_S^T H_\ell^T H_\ell X_S V_{S+i}^T F^{mT} H_k^T \,\middle|\, \mathcal{F}_{S+i-1}\right)$$

$$\text{(III.5)} \qquad = (X_S^T H_\ell^T H_\ell X_S) H_k F^m \mathbb{E}\left(V_{S+i} V_{S+i}^T \,\middle|\, \mathcal{F}_{S+i-1}\right) F^{mT} H_k^T$$

$$= (X_S^T H_\ell^T H_\ell X_S) H_k F^m Q_{S+i-1} F^{mT} H_k^T$$

$$< \|X_S\|^2 \times \text{constant} ,$$

since Q_t is assumed to be uniformly bounded.

As a consequence,

$$\text{(III.6)} \qquad \langle M,M\rangle_S \leqslant \text{constant} . A_S$$

Then, thanks to a suitable version of the law of large numbers for the martingales (see Appendix B), we get

$$\text{(III.7)} \qquad M_S/A_S \to 0 \text{ w.p.1 as } S \to \infty$$

since $A_S \to +\infty$ w.p.1 by assumption. Set

$$\text{(III.8)} \qquad P_S = \frac{1}{A_S} \sum_{t=1}^{S} X_t \, X_t^T \quad ;$$

(*) $\mathbb{E}(X/\mathcal{A})$ denotes the conditional expectation of X with respect to the σ - algebra $\mathcal{A}$.

from (III.7), (III.2) and (III.3), we get the following key result :

$$\text{(III.9)} \qquad \frac{1}{A_S}\,\Gamma_m(S) = HF^m P_S H^T + \varepsilon(S) ,$$

where the matrix $\varepsilon(S) \rightarrow 0$ ω.p.1 as $S \rightarrow \infty$. Finnaly, we get :

$$\text{(III.10)} \qquad \frac{1}{A_S}\,\mathcal{H}_{n,N}(S) = \begin{bmatrix} HP_S H^T & \cdots & HF^n P_S H^T & \cdots & HF^N P_S H^T \\ \vdots & & & & \vdots \\ HF^n P_S H^T & \cdots & HF^N P_S H^T & \cdots & HF^{N+n} P_S H^T \end{bmatrix} + \varepsilon(S) ,$$

with $\varepsilon(S) \rightarrow 0$ ω.p.1 as $S \rightarrow \infty$.

Step 2 : lower bounds for 1 , and end of the proof.

From (I.11), (I.17) and (III.10), we get

$$\text{(III.11)} \qquad [\hat{A}_n^N(S) - \hat{A}_n , \ldots , \hat{A}_1^N(S) - \hat{A}_1 , 0] \cdot \frac{1}{A_S^2}\,\mathcal{H}_{n,N}(S)\,\mathcal{H}_{n,N}^T(S) \rightarrow 0$$

ω.p.1 when $S \rightarrow \infty$, or, equivalently ,

$$\text{(III.12)} \qquad [\hat{A}_n^N(S) - \hat{A}_n , \ldots , \hat{A}_1^N(S) - \hat{A}_1] \cdot \frac{1}{A_S^2}\,\mathcal{H}_{n-1,N}(S)\,\mathcal{H}_{n-1,N}^T(S) \rightarrow 0$$

ω.p.1 when $S \rightarrow \infty$. Setting

$$\text{(III.13)} \qquad \mathcal{C}_N\,(F, P_S H^T) = (P_S H^T ,\ FP_S H^T , \ldots ,\ F^N\ P_S H^T) ,$$

we get with $\mathcal{O}_n$ as in (I.10).

$$\text{(III.14)} \qquad ([\hat{A}_n^N(S) - \hat{A}_n , \ldots , \hat{A}_1^N(S) - \hat{A}_1]\,\mathcal{O}_n)\,(\mathcal{C}_N(F,P_S H^T)\ \mathcal{C}_N^T(F,P_S H^T)\ \mathcal{O}_n^T) \rightarrow 0$$

ω.p.1 when $S \rightarrow \infty$. Then, for obtaining theorem 1, it remains to prove (cf. appendix A) that <u>the first</u> p (= dim X_t) <u>singular values of the matrix $\mathcal{C}_N\,\mathcal{C}_N^T\,\mathcal{O}_n$ of formula (III.14) are bounded from below ω.p.1 when $S \rightarrow \infty$.</u>

In appendix C, using a martingale argument, we prove the following lemma, where conditions (I.7) and (I.8) are used :

LEMMA : Setting

(III.15) $G_S G_S^T = \frac{1}{A_S} \sum_{t=0}^{S-1} V_{t+1} \; V_{t+1}^T$,

we have w.p. 1 when $S \to \infty$:

(III.16) $P_S - FP_S F^T - G_S G_S^T \to 0.$

Hence, the solution $\tilde{P}_S$ of the Lyapunov equation

(III.17) $\tilde{P}_S = F\tilde{P}_S F^T + G_S G_S^T$

satifies $\tilde{P}_S - P_S \to 0$ w.p.1 for $S \to \infty$. So we can replace P_S by $\tilde{P}_S$ in the formula (III.14), and we drop the superscript " $\tilde{}$ " for simplicity.

Now, the bracket (H,F,G_S) define a stationary Gauss-Markov process of the form (II.1), and we shall apply the results of section II to this process. But, with the rotations (II.8), we get, using appendix A :

(III.18) $\sigma_{min}(\mathbb{C}_N(F,P_S H^T)\, \mathbb{C}_N^T(F,P_S H^T)\mathbb{O}_n^T) > \sigma_{min}^2(\mathbb{C}_N(F,P_S H^T)\; \sigma_{min}(\mathbb{O}_n^T).$

On the other hand, one verifies that

(III.19) $\mathbb{C}_\infty(F,P_S H^T) = \mathbb{C}_\infty(F,G_S)\; \mathbb{R}^{T/2}(S)$,

where $\mathbb{R}(S)$ is the infinite covariance matrix of the stationary process $y_t(S)$ generated by the bracket (H,F,G_S). We shall finish the proof by giving

- lower bounds for $\sigma_{min}\ (\mathbb{C}_\infty(F,G_S))$
- lower bounds for $S_{min}\ (Y_t^S)$
- upper bounds for $\mathbb{C}_\infty(F,P_SH^T)\ \mathbb{C}_\infty^T(F,P_SH^T) - \mathbb{C}_N(F,P_SH^T)\ \mathbb{C}_N^T(F,P_SH^T)$,

and using (III.18) together with theorem 4.

<u>Lower bounds for σ_{min}</u> $(\mathbb{C}_\infty(F,G_S))$.

Here, we use simply assumption (I.7) $G_SG_S^T > G_*G_*^T$, which implies, since the pair (F,G_*) is controllable :

$$\text{(III.20)}\quad \mathbb{C}_\infty(F,G_S)\ \mathbb{C}_\infty^T(F,G_S) \geqslant \mathbb{C}_\infty(F,G_*)\ \mathbb{C}_\infty^T(F,G_*) > 0 ,$$

hence the corresponding inequality for the singular values.

<u>Lower bounds for S_{min}</u>(Y_t^S) .

Here, again, we have, for $0 \leqslant \theta < 2\pi$

$$\text{(III.21)}\quad \begin{aligned} &H(I - e^{i\theta}F)^{-1}\ G_SG_S^T\ (I - e^{-i\theta}F^T)^{-1}\ H^T \\ &\geqslant H(I - e^{i\theta}F)^{-1}\ G_*G_*^T\ (I - e^{i\theta}F^T)^{-1}\ H^T , \end{aligned}$$

so that, with Y_t^* generated by the bracket (H,F,G_*) :

$$\text{(III.22)}\quad S_{min}(Y_t^S) \geqslant S_{min}(Y_t^*) > 0 .$$

<u>Upper bounds for</u> $\mathbb{C}_\infty(F,P_SH^T)\ \mathbb{C}_\infty^T(F,P_SH^T) - \mathbb{C}_N^T(F,P_SH^T)\ \mathbb{C}_N^T(F,P_SH^T)$.

Since $A_S = Tr \sum_1^S X_tX_t^T$, P_S is bounded uniformly in S. Hence there exists a constant c such that $P_SH^THP_S < c\ I$, so that

$$\text{(III.23)}\quad \begin{aligned} &\mathbb{C}_\infty(F,P_SH^T)\ \mathbb{C}_\infty^T(F,P_SH^T) - \mathbb{C}_N(F,P_SH^T)\ \mathbb{C}_N^T(F,P_SH^T) \\ &= \sum_{n<N} F^n(P_SH^THP_S)\ F^{n^T} \\ &\leqslant \text{constant} \times (\sum_{n>N} \rho^{2n})\ .\ I\ , \end{aligned}$$

where ρ denotes the spectral radius of F. Since F is asymptotically stable, we have $\rho < 1$. Hence the first member of (III.23) tends to zéro uniformly in S when $N \to \infty$.

Finally, thanks to (II. 10-b) of theorem 4, and using (III.23),(III.20) and (III.22), we obtain that there exists N such that

$$\text{(III.24)} \quad \liminf_{S \to \infty} \sigma_{min}(\mathbb{C}_N(F,P_S H^T)) > 0 ,$$

which finishes the proof of theorem 1, thanks to (III.14) and (III.18).

The proof of theroem 2 goes through the same steps as above. Here, the key point is that (III.14) is now replaced by

$$\text{(III.25)} \quad \begin{aligned} &\frac{1}{A_S^2} (\hat{F}^N(S) - \hat{F}) \overset{HO}{\mathbb{H}_N}(S) \overset{HO}{\mathbb{H}_N}(S)^T \to 0 \; ; \\ &\frac{1}{A_S^2} (\hat{H}^N(S) - \hat{H}) \overset{HO}{\mathbb{H}_N}(S) \overset{HO}{\mathbb{H}_N}(S)^T \to 0 \end{aligned}$$

ω.p.1 when $S \to \infty$; but

$$\text{(III.26)} \quad \frac{1}{A_S^2} \overset{HO}{\mathbb{H}_N}(S) \overset{HO}{\mathbb{H}_N}(S) = \overset{HO}{\mathbb{O}} \; \mathbb{C}_N(F,P_S H^T) \; \mathbb{C}_N^T(F,P_S H^T) (\overset{HO}{\mathbb{O}})^T + \varepsilon(S) ,$$

so that we can proceed as before, and obtain the theorem 2, taking into account the fact that $\overset{HO}{\mathbb{O}}$ is here a square invertible matrix.

The proof of theorem 3 is also similar. Starting with $\Gamma_1(S)$ instead of $\Gamma_0(S)$ is the Hankel matrix allows us to annihilate the effect of the measurement noise, using again a martingale argument. Here, the relevant factorization is

$$\text{(III.27)} \quad \frac{1}{A_S^2} \mathbb{H}_{n,N}(S) \mathbb{H}_{n,N}^T(S) = \mathbb{O}_n \; \mathbb{C}_N(F,FP_S H^T) \; \mathbb{C}_N^T(F,FP_S H^T) \; \mathbb{O}_n^T + \varepsilon(S),$$

and we finish the proof using the lower bounds in (II.26).

Appendix A : Some results on the Singular Values of a matrix.

Recall that, given a mxp matrix A, we have the Singular Value Decomposition (SVD) of A ([24])

(A.1) $$A = U_m \Delta_{m,p} V_p^T$$

where U_m and V_p^T are orthogonal matrices, and the mxp matrix $\Delta_{m,p}$ is of the form (here we assume $m < p$, for example).

(A.2) $$\Delta_{m,p} = \begin{bmatrix} \sigma_1 & & & \\ & \ddots & & 0 \\ 0 & & \sigma_m & \end{bmatrix}$$

where the singular values of A are the ordered real $\sigma_1 \geqslant \ldots \geqslant \sigma_m \geqslant 0$. We shall use the following facts :

(A.3) $$\|A\| \triangleq \sup_{\|x\| = 1} \|Ax\| = \sigma_1 \text{ , denoted by } \sigma_{max}(A)$$

(A.4) $$\sigma_{max}(AB) \leqslant \sigma_{max}(A)\ \sigma_{max}(B)$$

(A.5) $$\sigma_{max}(A) = (\lambda_{max}(AA^T))^{1/2} = (\lambda_{max}(A^TA))^{1/2} ,$$

where λ_{max} denotes the greatest eigenvalue.

The corresponding results for the lowest singular values are some what more involved.

Let us denote by $\sigma_{min}(A)$ the lowest singular value σ_m in (A.2). Then we have the following result

(A.6) $$\sigma_{min}(AB) \geqslant \sigma_{min}(A)\ \sigma_{min}(B) \quad \underline{\text{if B is invertible}}.$$

PROOF : For $A = U_A \Delta_A V_A^T$ (as in (A.1)), set

$$(A.7) \qquad A^{\dagger} = V_A \Delta_A^{-T} U_A^T ,$$

where, according to the notations of (A.2)

$$(A.8) \qquad \Delta_A^{-T} = \begin{bmatrix} 1/\sigma_1 & & 0 \\ & \ddots & \\ & & 1/\sigma_m \\ & 0 & \end{bmatrix} , \text{ with } 1/0 \triangleq 0.$$

Now assume A is nxp matrix and B is pxp with $p \geqslant n$. Then

$$(A.9) \qquad ABB^{\dagger}A^{\dagger} = ABB^{-1}A^{\dagger} = U_A \Delta_A \Delta_A^{-T} U_A^{-T} = U_A \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix} U_A^T ,$$

where $r \leqslant n$ is the rank of A. If $r < n$, (A.6) reduces to $0 = 0$; if $r = n$, using (A.9) and (A.3), and thanks to $AB\, x = AB\,(B^{\dagger}A^{\dagger}AB\, x)$ for every x, we get

$$(A.10) \qquad \sigma_{\min}(AB) \triangleq \min_x \frac{\|AB\, x\|}{\|x\|} \geqslant \min_y \frac{\|ABB^{\dagger}A^{\dagger}y\|}{\|B^{\dagger}A^{\dagger}y\|} = 1/\sigma_{\max}(B^{\dagger}A^{\dagger}),$$

hence (A.6), using (A.4). When $p \leqslant n$, apply the argument to B^TA^T, the key point being $\Delta_A^T \Delta_A^{-1} = I_p$ when A is of full rank. (A.6) is <u>false</u> in the general case ∎

Now <u>assume</u> A <u>is nxp</u>, B <u>is pxm with $p = \min(m,n,p)$</u>. Set $\sigma_{\min}(AB) \triangleq \sigma_p(AB)$, which is the lowest non trivial singular value of AB (rank (AB) $\leqslant$ p). Then, again

$$(A.11) \qquad \sigma_{\min}(AB)\ (\triangleq \sigma_p(AB)) \geqslant \sigma_{\min}(A)\ \sigma_{\min}(B)$$

PROOF : Here again we can assume A and B of full rank. Then, using $\Delta_B \Delta_B^{-T} = I_p$, we get

$$\text{(A.12)} \qquad ABB^{\dagger}A^{\dagger} = U_A \, \Delta_A \, \Delta_A^{-T} \, U_A^T = U_A \begin{bmatrix} I_p & 0 \\ 0 & 0 \end{bmatrix} U_A^T$$

so that, again ,

$$\text{(A.13)} \qquad \begin{aligned} \sigma_{min}(AB) \;\triangleq\; \sigma_p(AB) &= \min \frac{\|ABB^{\dagger}A^{\dagger}y\|}{\|B^{\dagger}A^{\dagger}y\|} \\ &\geq \frac{1}{\sigma_{max}(B^T A^T)} \quad , \end{aligned}$$

hence (A.11) ∎

The inequalities (A.6) and (A.11) are used in the section II for obtaining the lower bounds on the singular values of the Hankel matrices. The lower bound can't be directly obtained for $\mathcal{C}_N(F,P_S H^T)$ since this matrix is the product of two matrices A and B with $p > n$ and m, and neither (A.11) nor (A.6) can be applied in this case.

APPENDIX B : A strong law of large numbers of martingales.

The following result is proved in [25] :

LEMMA : *Let M_t be a locally square integrable martingale with respect to a family (F_t) of σ - algebras, such that $M_o = 0$. Set*

(B.1) $$\langle M,M\rangle_t = \sum_{s=1}^{t} \mathbb{E}((M_s - M_{s-1})^2 / F_{s-1}).$$

Then, w.p.1 we have :

(B.2) $$\frac{1}{\langle M,M\rangle_t} M_t \to 0 \text{ on the set } \left\{\lim_{t\to\infty} \langle M,M\rangle_t = +\infty\right\},$$

(B.3) $$\lim_{t\to\infty} M_t \text{ exists and is finite on the set } \left\{\lim_{t\to\infty} \langle M,M\rangle_t < +\infty\right\}.$$

An immediate consequence is that the lemma still holds if $\langle M,M\rangle_t$ is replaced by a process A_t satisfying

(B.4) $$\langle M,M\rangle_t \leq A_t \quad \text{w.p.1.}$$

This is the results we use in the martingale arguments.

APPENDIX C : miscellaneous.

C.1 : PROOF OF (III.16)

We have

$$
\begin{aligned}
\sum_0^S X_t X_t^T &= X_0 X_0^T + \sum_0^{S-1} (FX_t + V_{t+1})(FX_t + V_{t+1})^T \\
&= X_0 X_0^T + F(\sum_0^{S-1} X_t X_t^T) F^T + \sum_1^S V_t V_t^T \\
&\quad + F(\sum_0^{S-1} X_t V_{t+1}^T) + (\sum_0^{S-1} V_{t+1} X_t^T) F^T \\
&= (1) + (2) + (3) + (4) + (5) .
\end{aligned}
\tag{C.1}
$$

Neglecting $X_0 X_0^T$ and $X_S X_S^T$ with respect to A_S, we have only to prove that (4) and (5) are also neglectible with respect to A_S. This is obtained by applying the argument used in (III.3) to (III.5). Note we have only used the conditions : $A_S \uparrow + \infty$, and Q_t is uniformly bounded.

C.2 : PROOF OF LEMMA 1

Step 1 :

$$\frac{1}{S} \sum_1^S V_t V_t^T \geq G_* G_*^T - \varepsilon(S).I$$

with $\varepsilon(S) \to 0$ w.P.1 as $S \to \infty$.

PROOF : Under (I.9), we can decompose the noise V_t into

$$V_t = \overset{*}{V}_t + \tilde{V}_t \ , \quad \mathrm{Cov}\ \overset{*}{V}_t = G_* G_*^T , \quad \overset{*}{V} \text{ and } \tilde{V} \text{ orthogonal.} \tag{C.2}$$

Then we get

$$(C.3)\qquad \sum_1^S V_t\, V_t^T = \sum_1^S \overset{*}{V}_t\, \overset{*}{V}{}_t^T + \sum_1^S \tilde{V}_t\, \tilde{V}_t^T + \sum_1^S (\overset{*}{V}_t\, \tilde{V}_t^T + \tilde{V}_t\, \overset{*}{V}{}_t^T).$$

Consider the scalar martingale (k,ℓ fixed)

$$(C.4)\qquad M_S = \sum_1^S \overset{*}{V}{}_t^{\ell}\, \tilde{V}_t^k \; ;$$

we have

$$(C.5)\qquad \langle M,M\rangle_S = Q_*^{\ell\ell}\Big(\sum_1^S \tilde{Q}_t^{kk}\Big), \quad Q_* = G_* G_*^T, \quad \tilde{Q}_t = Q_t - Q_* .$$

Since $\tilde{Q}_t$ is bounded, we get, using appendix B, $\frac{1}{S} M_S \to 0$, hence the result.

<u>Step 2 : $A_S \uparrow + \infty$ w.p.1</u>

<u>PROOF</u> : Assume $\mathbb{P}\{\lim A_S < +\infty\} > 0$. Then, considering the martingale

$$M_S = \sum_0^{S-1} X_t^{\ell}\, V_{t+1}^k$$

for k and ℓ fixed, we can use the remark in the appendix C for obtaining that, on the set $\{\lim A_S < +\infty\}$, $\sum_0^{S-1} X_t\, V_{t+1}^T$ converges to a finite value w.p.1. Then, taking the trace on both sides of (C.1) we get

$$(C.6)\qquad \lim_{S\to\infty} \sum_1^S \|V_t\|^2 \text{ exists and is finite w.p.1 on the set}$$

$\{\lim A_S < \infty\}$, which is impossible, in view of the results of step1.

Step 3 :

$$0 < \liminf_{S \to \infty} \frac{A_S}{S} \leqslant \limsup_{S \to \infty} \frac{A_S}{S} < +\infty \qquad \omega.p.1$$

PROOF : Since $A_S \uparrow +\infty$ w.p.1 , we can neglect the terms 1 , 4 and 5 in (C.1), thus obtaining

$$\text{(C.7)} \qquad \frac{1}{S} \sum_{0}^{S-1} X_t X_t^T = F\left(\frac{1}{S} \sum_{0}^{S-1} X_t X_t^T\right) F^T + \frac{1}{S} \sum_{1}^{S} V_t V_t^T + \varepsilon(S) ,$$

hence, thanks to the result of step 1, $\frac{1}{S} \sum_{0}^{S-1} X_t X_t^T$ is uniformly bounded from above and from below, which implies obviously the result of step 3.

The lemma 1 follows then obviously.

CONCLUSION AND DISCUSSION

It has been shown that the classical linear methods (Instrumental Variable method, Ho-Kalman algorithm,...) for identifying the pole part of a Gauss-Markov model are robust in the presence of nonstationary excitation noise, and that the identification can be achieved using only a single sample of the corresponding (nonstationary) signal.

The accuracy of these identification methods may be evaluated by looking at the bounds on the Hankel matrices which are given in the section II : it is clear that well excited modes are identified with more accuracy than the other ones.

A question remains open : what happens when model reduction is done in the nonstationary case ? Note that the martingale argument, the proof is based upon, no longer holds in the case of underparametrization ; hence, obtaining consistency results in the nonstationary case with model reduction seems to be a hard task. Nevertheless, the experimental results reported in [1][2][3] show that the methods still work in reasonable situations.

REFERENCES.

[1] M. PREVOSTO : Thèse de Docteur Ingénieur, IRISA, Rennes, 1982.

[2] M. PREVOSTO, A. BENVENISTE, B. BARNOUIN : "Identification of Vibrating structures subject to Nonstationary Excitation : a non Stationary Stochastic Realization Problem" ICASSP — 1982, Paris.

[3] M. PREVOSTO, A. BENVENISTE, B. BARNOUIN : "Modelling and Identification of vibrating structures : a Large Scale, Non-stationary, Stochastic Realization Problem" IRISA Report 1982, submitted to IEEE - AC.

[4] R.L. GOODRICH, P.E. CAINES : "On the Nonstationary Covariance Realization Problem", IEEE AC - 24 n°5, 1979.

[5] R.L. GOODRICH, P.E. CAINES : "Linear System Identification from Nonstationary Cross-Sectional Data" IEEE AC - 24 n°3, 1979.

[6] Y. BARAM : "Realization and Reduction of Markovian Models from Nonstationary Data "IEEE AC - 26 n°6, 1981.

[7] K.Y. WONG, E. POLAK : "Identification of Linear Discrete Time Systems Using the Instrumental Variable Method" IEEE AC - 12 n°6, 1967.

[8] T. SÖDERSTRÖM, P. STOICA : "Comparison of Some Instrumental Variable Methods : Consistency and Accuracy Aspects" Automatica Vol. 17 n°1, 1981.

[9] J. RISSAMEN, P.E. CAINES : "The Strong Consistency of Maximum Likelihood Estimators for ARMA Processes", Annals of Stat. Vol. 7 n°2, 1979.

[10] P.E. CAINES : "On the Asymptotic Normality of Instrumental Variable and Least Squares Estimators" IEEE AC August 1976.

[11] M. PREVOSTO, A.BENVENISTE, B. BARNOUIN : "La méthode des Variables Instrumentales en treillis ; application à la modélisation des structures en vibration", Aussois Worshop on "Fast Algorithms for Linear Systems", 21-25 Sept 1981, édited in "Outils et Modèles Mathématiques pour l'analyse des systèmes, l'automatique, et le traitement du Signal" Vol. 2, 1982, Editions du CNRS, Paris.

[12] B.L. HO, R.E. KALMAN : "Effective construction of linear state-variable models from input/output functions " Proc. 3rd Allerton Conf. Circuit and Systems Theory, pp. 449-459, 1965.

[13] J. RISSANEN : "Recursive Identification of Linear Systems" SIAM J. Control 9, 420-430, 1971.

[14] L.M. SILVERMAN : "Realization of Linear Dynamical Systems" IEEE AC - 16, 1971.

[15] G. RUCKEBUSCH : "Représentations Markoviennes de Processus Gaussiens Stationnaires" Thèse de 3è cycle, Paris VI, 1975.

[16] P. FAURRE : "Réalisations markoviennes de processus stationnaires" Rapport IRIA n°13, 1973.

[17] A. LINDQUIST, G. PICCI : "On the Stochastic Realization Problem" SIAM J. Cont . Optimiz. vol. 17, May 1979.

[18] J. STAAR : "A study of MBH - type Realization Algorithms" Automatica Vol. 17 n°3, 1981.

[19] Y.V. GENIN, S. KUNG : "Rational Approximation with Hankel Norm criterion" Proc. 1980 IEEE - CDC, New Mexico.

[20] Y.V. GENIN, S. KUNG : "A two - Variable Approach to the Model Reduction Problem with Hankel Norm Criterion" IEEE CAS - 28 n°9, 1981.

[21] L.M. SILVERMAN, M. BETTAYEB : "Optimal Approximation of Linear Systems" IEEE AC - Déc. 1980.

[22] ROZANOV : "Stationary Random Processes" Holden Day, 1967.

[23] H. DYM, H.P. Mc KEAN : "Gaussian Processes Function Theory, and the Inverse Spectral Problem" Academic Press 1976.

[24] V.C. KLEMA, A.J. LAUB : " The Singular Value Decomposition : its computation and Some Applications II ; IEEE AC - 25 n°2, 1980.

[25] J. NEVEU : "Discrete Parameter Martingales" North Holland, 1975.

Liste des Publications Internes IRISA

PI 130 **Un système d'enchères simultanées réparti**
M. Banâtre, A.Couvert , 27 pages ; *Janvier 1980*

PI 131 **Méthodes Pseudo-Booléennes**
G. Boulaye , 14 pages ; *Janvier 1980*

PI 132 **Semaine d'étude internationale sur la détection et l'estimation temps-réel des contours et des mouvements dans les images**
Organisation : CCETT, INSA, IRISA (Rennes), , 260 pages ; *Septembre 1979*

PI 133 **Weak solutions for semi-martingales**
J. Pellaumail , 19 pages ; *Mars 1980*

PI 134 **Processus non sequentiel et leurs observations en univers non-centralisé**
Ph. Darondeau , 53 pages ; *Avril 1980*

PI 135 **Synchronisation and protection features for data abstraction**
D. Herman, M.Raynal , 21 pages ; *Juin 1980*

PI 136 **Elaboration et évaluation d'un graphe d'implication pour des données binaires**
I.C. Lerman, R. Gras, H. Rostam , 70 pages ; *Août 1980*

PI 137 **Rapport à la C.E.E. sur les techniques de programmation**
J. André, D. Coan, H. Geist, D. Law, Y. Letertre, H. Schoenen , 96 pages ; *Septembre 1980*

PI 138 **Définition d'un logiciel de transport adapté à la mise en oeuvre d'application burotiques sur des réseaux locaux**
S. Gaucher Cazalis, F. Krier, H. Le Goff , 70 pages ; *Septembre 1980*

PI 139 **Efficacité des algorithmes récursifs en présence de systèmes non-stationnaires.**
A. Benveniste, G. Ruget , 35 pages ; *Août 1980*

PI 140 **Structures de communication extensibles**
P. Le Guernic, M. Raynal , 60 pages ; *Octobre 1980*

PI 141 **Comparaison de tableaux de fréquence**
B. Escoffier , 16 pages ; *Octobre 1980*

PI 142 **Un lemme général de stabilité pour la commande adaptative en déterministe de systèmes non nécessairement à minimum de phase]:**
Cl. Samson , 40 pages ; *Novembre 1980*

PI 143 **Détection, Estimation de l'orientation et saisie d'une cible mobile par proximétrie optique**
B. Espiau , 142 pages ; *Janvier 1981*

PI 144 **Une contribution à l'étude de l'impact de l'informatique sur les organisations**
L. Breton, A. Prod'homme, J. Villard , 58 pages ; *Décembre 1980*

PI 145 **Rupture de modèles statistiques**
M. Basseville, A. Benveniste , 130 pages ; *Mars 1981*

PI 146 **Traitement des questionnaires avec non réponse, analyse des correspondances avec marge modifiée et analyse multicanonique avec contrainte**
B. Escoffier , 38 pages ; *Mars 1981*

PI 147 **Deux files d'attente à capacité limitée en tandem**
J. Pellaumail, J. Boyer , 19 pages ; *Juillet 1981*

PI 148 **Programme de classification hiérarchique : 1) Méthode de la vraisemblance des liens, 2) Méthode de la variance expliquée**
I.C. Lerman , 113 pages ; *Juin 1981*

PI 149 **Convergence des méthodes de commande adaptative en présence de perturbations aléatoires**
J.J. Fuchs , 46 pages ; *Juillet 1981*

PI 150 **Construction automatique et évaluation d'un graphe d'«implication» issu de données binaires, dans le cadre de la didactique des mathématiques**
H. Rostam , 112 pages ; *Juin 1981*

PI 151 **Réalisation d'un outil d'évaluation de mécanismes de détection de pannes[-]Projet Pilote SURF**
B. Decouty, G. Michel, C. Wagner, Y. Crouzet , 59 pages ; *Juillet 1981*

PI 152 **Règle maximale**
J. Pellaumail , 18 pages ; *Septembre 1981*

PI 153 **Corrélation partielle dans le cas « qualitatif »**
I.C. Lerman , 125 pages ; *Octobre 1981*

PI 154 **Stability analysis of adaptively controlled not-necessarily minimum phase systems with disturbances**
Cl. Samson , 40 pages ; *Octobre 1981*

PI 155 **Analyses d'opinions d'instituteurs à l'égard de l'appropriation des nombres naturels par les élèves de cycle préparatoire**
R. Gras , 37 pages ; *Octobre 1981*

PI 156 **Récursion induction principle revisited**
G. Boudol, L. Kott , 49 pages ; *Décembre 1981*

PI 157 **Loi d'une variable aléatoire à valeur R^+** réalisant le minimum des moments d'ordre supérieur à deux lorsque les deux premiers sont fixés
M.Kowalowka, R. Marie , 8 pages ; *Décembre 1981*

PI 158 **Réalisations stochastiques de signaux non stationnaires, et identification sur un seul échantillon**
A. Benveniste J.J. Fuchs , 33 pages ; *Mars 1982*

PI 159 **Méthode d'interprétation d'une classification hiérarchique d'attributs-modalites pour l'«explication» d'une variable ; application à la recherche de seuil critique de la tension artérielle systolique et des indicateurs de risque cardiovasculaire**
B. Tallur , 34 pages ; *Janvier 1982*

PI 160 **Probabilite stationnaire d'un réseau de files d'attente multiclasse à serveur central et à routages dépendant de l'état**
L.M. Le Ny , 18 pages ; *Janvier 1982*

PI 161 **Détection séquentielle de changements brusques des caractéristiques spectrales d'un signal numérique**
M. Basseville, A. Benveniste , pages ; *Mars 1982*

PI 162 **Actes regroupés des Journées de Classification de Toulouse (Mai 1980), et de Nancy (Juin 1981)**
I.C. Lerman , 304 pages ; *Mars 1982*